

一种新的改进遗传算法及其性能分析*

吴养会¹, 王乃信¹, 王正中²

(1 西北农林科技大学 生命科学学院, 陕西 杨凌 712100;

2 西北农林科技大学 水利与建筑工程学院, 陕西 杨凌 712100)

[摘要] 针对简单遗传算法收敛中所存在的收敛速度慢及局部收敛问题, 引入了一种新的改进遗传算法。该算法利用不断淘汰相似个体, 并不断补充新个体的方法增加种群的多样性。并用一个复杂的函数对算法进行测试, 结果表明该算法性能优于简单遗传算法。

[关键词] 遗传算法; 淘汰相似个体; 改进算法

[中图分类号] O 224

[文献标识码] A

[文章编号] 1671-9387(2004)09-0124-03

遗传算法是由美国Michigan大学的Holland教授根据生物进化理论和遗传变异理论提出的一种基于种群搜索的优化算法。其思想是随机产生初始种群, 通过选择(reproduction)、交叉(crossover)和变异(mutation)等遗传算子的共同作用使种群不断进化, 最终得到最优解。与其他优化方法相比, 遗传算法以单一字符串的形式描述所研究的问题, 只需利用适应度函数来进行优化计算, 而不需要函数导数等其他信息, 特别适合解决其他科学技术无法解决或难以解决的复杂问题, 如结构优化、非线性优化、机器学习等, 是继专家系统、人工神经网络之后又一受人青睐的人工智能学科^[1]。

遗传算法在发展初期称为简单遗传算法(Simple GA SGA), 简单遗传算法思路直观、操作简单, 但其收敛速度慢, 更重要的是简单遗传算法不能保证收敛到全局最优解。现有的大多数改进遗传算法大都局限于 p_c, p_m (交叉、变异概率)的改进, 忽略了算法整体设计^[2,3]。本文提出了一种新的改进遗传算法(Improved GA, IGA), 其思想是在自适应 p_c, p_m 选择的基础上, 在算法执行中采用淘汰相似个体的方法, 并不断补充父代优良个体, 提高种群多样性, 加快算法的收敛速度及收敛到全局最优解的能力。

1 一种改进的遗传算法

首先说明各量的意义: l 指 l 个体, 即长度为 l 的 0 和 1 字符串; β^l 表示齐次种群的全体 (对于种群

X , 若任意 $i, j = 1, 2, \dots, N, X_i = X_j$ 恒成立, 则称 X 为齐次种群); $o(h)$ 代表模式 H 的阶 (模式 H 中确定位置的个数); T 为总进化代数; f 表示适应值; f_{\max} 表示最大适应值; f_{avg} 为种群的平均适应值。

相同长度的以 a 为基的两个字符串中对对应不相同的数量称为两者间的广义海明距离。记 $X(t)$ 为第 t 代进化种群, 其第 i 个和第 j 个个体分别为 $X_i(t)$ 和 $X_j(t)$, 它们的广义海明距离为

$$H(X_i(t), X_j(t)) = \sum_{k=1}^l |x_{ik}(t) - x_{jk}(t)|$$

如含有相同基 $1 * * 1 0 0 1 * * * * 1$ 的两个个体

$$X_i(t) = 1 0 0 1 0 0 1 1 1 0 0 1 1,$$

$$X_j(t) = 1 1 0 1 0 0 1 0 1 1 0 0 1.$$

它们间的广义海明距离为

$$H(X_i(t), X_j(t)) = \sum_{k=1}^{13} |x_{ik}(t) - x_{jk}(t)| = 4$$

1.1 改进算法

在遗传算法中交换是产生新个体的主要手段, 它类似于生物学的杂交, 使不同个体的基因互相交换, 从而产生新个体, 扩大搜索空间。但当进化代数接近终止代数时, 种群接近于一个齐次种群。若 X 接近于一个齐次种群, 其个体间的交叉操作便很难产生新的个体, 于是这种母体通过交叉操作产生新的子代个体的进化方式成为近亲繁殖。近亲繁殖不利于产生多样性种群。这种未成熟收敛是遗传算法中

* [收稿日期] 2004-02-16

[作者简介] 吴养会(1971-), 男, 陕西乾县人, 在读硕士, 主要从事农业工程中的数学方法研究。

不可忽视的现象,其结果是导致算法不能收敛于全局最优解,影响算法的收敛性能。

本文在文献[3]的基础上对GA的求解提出如下的改进措施与步骤:

(1)淘汰相似个体。在算法执行过程中,为减少齐次种群出现的机会,对个体按适应度值大小进行排序,依次计算相邻个体间的广义海明距离 H 。如果广义海明距离小于 d ,就删除其中适应度较小的个体,代之以父代群体中的优秀个体,否则,保留这两个个体加入到中间群体。对 d 的数值应当适当选择,以提高群体的多样性, d 值过大可能将优秀个体删除,过小则可能起不到过滤作用,其数值应随进化代数而变化。在进化过程的前期,增加种群的多样性是主要目的,此时应使 d 值大一些,而到进化的后期,种群趋于齐次种群,增加种群的多样性已不是主要目的,此时 d 的选取应很小。因而 d 应是变量,且随进化代数的增加而减少,最终趋于零,这是符合进化规律的。基于上述原因,参数 d 可取为

$$d = k(l - o(h)) \left(1 - \frac{t}{T}\right), (k < 1.0)。$$

(2)自适应交叉和变异操作。对于中间群体,依文献[4]所提出的 p_c, p_m 的确定方法进行交叉、变异操作,产生下一代群体。其中

$$p_c = k_1(f_0 - f_{\max}) / (f_0 - f_{\text{avg}}),$$

$$p_m = k_2(f_0 - f) / (f_0 - f_{\text{avg}}), (k_1, k_2 = 1.0)。$$

在保证高适应值的解加快GA收敛的同时,防止低适应值的解使遗传算法陷入局部最小。

(3)父代最优个体保留操作。交叉和变异操作执行后,当父代种群最优个体的适应度大于子代最劣个体的适应度时,以前者取代后者,从而使子代种群的性能不劣于父代种群。这种思想也体现了优胜劣汰的生物进化规律^[5]。

1.2 算法设计

(1)确定种群规模 N ,选择交叉和变异的方法分别按适应度比例选择法和自适应法,算法终止条件是进化代数等于进化终止条件。

(2)以二进制编码,随机产生初始种群 $X(0)$ 。

(3)对第 t 代群体按适应度进行排序,依次计算广义海明距离。

(4)判断广义海明距离数值和参数 d 的大小关系,若 $H < d$,依次用父代中的个体替换适应度小的个体。否则,执行下一步。

(5)对得到的中间种群,采用自适应 p_c, p_m 进行交叉、变异操作。

(6)终止条件判断。若不满足终止条件,令 $t = t + 1$,转到(3),执行下一步遗传操作,否则输出当前最优个体,算法结束。

2 算法的具体实现

为验证算法的有效性,就算法的全局收敛性和收敛速度与简单遗传算法进行比较。考虑具有相当复杂度的测试函数Shaffer F_6 函数^[6]:

$$f(x_1, x_2) = 0.5 - \frac{\sin^2 \sqrt{x_1^2 + x_2^2} - 0.5}{(1 + 0.001(x_1^2 + x_2^2))^2},$$

$$\text{S. T. } -100 < x_i < 100,$$

2.1 算法的实现步骤

该函数有无限个局部极大点,其中只有一个(0, 0)为全局极大,最大值为1。此函数最大值峰周围有一个圈脊,其取值均为0.990283,因此很容易停滞在此局部极大点。解决该问题的构造步骤如下^[7,8]:

(1)确定决策变量和约束条件。 $-100 < x_i < 100, i = 1, 2$ 。

(2)确定编码方案。用长度为11的二进制编码串表示2个变量 x_1 和 x_2 。再将二者的2个11位长的二进制编码串连接在一起,组成一个22位长的二进制编码,即构成了Shaffer F_6 函数:优化问题的染色体编码方法。

(3)确定解码方法。解码时需要将22位长的二进制编码串切断为2个11位的二进制编码串,然后将其转换成对应的十进制整数代码,记为 y_1 和 y_2 。将 y_i 转换为 x_i 的解码公式为

$$x_i = \frac{200y_i}{2047} - 100, i = 1, 2。$$

(4)确定个体评价方法。由于Shaffer F_6 函数的值总是非负值,并且优化目标是求函数的最大值,所以将个体的适应度函数直接取为对应的目标函数值。

(5)关于遗传算子。选择运算使用比例选择算子,交叉和变异运算使用自适应 p_c, p_m 确定的算子进行。

(6)确定遗传算法运行参数。群体大小 $M = 100$,终止代数 $T = 200$ 。

2.2 计算结果

表1为简单遗传算法和改进遗传算法的收敛结果比较,分别比较迭代50, 100, 150和200次时的函数值。从表1可以看出,对于简单遗传算法,经过几百次的进化,其解分别收敛于局部极大值,而始终没

有收敛到全局最大点(0, 0)。而对于改进遗传算法, 均收敛于全局最大值。进化经过几百代甚至几十代后, 无论初始群体如何,

表1 两种算法收敛比较

Table 1 Comparison on the convergence of the two algorithms

算法 Algorithm	进化代数 T The number of evolutionary generations			
	50	100	150	200
SGA	0.889 4	0.875 6	0.885 5	0.980 3
IGA	0.972 5	0.973 2	0.982 4	0.999 2

3 结 论

为了避免种群过早收敛于齐次种群, 本文提出了一种改进遗传算法, 在进化过程中不断淘汰相似个体, 最大限度减少齐次种群中近亲繁殖的机会, 并

不断补充新个体, 从而避免了搜索空间的迅速缩小。再利用自适应选择 p_c, p_m 进行交叉和变异, 大大提高了全局搜索性能和收敛速度。实例验证结果表明, 这种改进算法的性能明显优于简单遗传算法。由此可见, 该算法是一种较优的搜索算法。

[参考文献]

- [1] Holland J H. A daptation in N ational and A rtificial System s[M]. M ichigan: The U niversity of M ichigan Press, 1975.
- [2] 巩敦卫, 孙小燕, 郭西进. 一种新的优胜劣汰遗传算法[J]. 控制与决策, 2002, 11(6): 908- 912.
- [3] 韩万林. 遗传算法的改进[J]. 中国矿业大学学报, 2001, 3(1): 102- 105.
- [4] 陈长征, 王 楠. GA 中交叉和变异概率选择的自适应方法及作用机理[J]. 控制理论与应用, 2002, 2(1): 41- 44.
- [5] 张文修, 梁 怡. 遗传算法的数学基础[M]. 西安: 西安交通大学出版社, 2000.
- [6] 云庆夏. 进化算法[M]. 北京: 冶金工业出版社, 2000.
- [7] 王小平, 曹立明. 遗传算法理论、应用与软件实现[M]. 西安: 西安交通大学出版社, 2002.
- [8] 张 晖, 吴 斌. 引入模拟退火机制的新型遗传算法[J]. 电子科技大学学报, 2003, 2(1): 39- 42.

A new improved genetic algorithm and its property analysis

WU Yang-hui¹, WANG Nai-xin¹, WANG Zheng-zhong²

(¹ College of Life Sciences, Northwest Sci-Tech University of Agriculture and Forestry, Yangling, Shaanxi 712100, China;

² College of Water Resources and Architectural Engineering, Northwest Sci-Tech University of Agriculture and Forestry, Yangling, Shaanxi 712100, China)

Abstract: In order to avoid the slow-convergence and local convergence of simple genetic algorithm (SGA), a kind of new improved genetic algorithm (IGA) was proposed in this paper. This algorithm used unceasing elimination of similar individual method to increase the multiplicity of population, and has been checked with a complex function. The experimental results show IGA has great advantage over SGA.

Key words: genetic algorithm; unceasing elimination of similar individual; IGA